Sherry [illegible]rrell, Ph.D.
NISS/[illegible]/Ord/HCFA
6325 Security Blvd. 2B-[illegible]OM
Baltimore, MD 21207
(301) 597-1416 FTS 597-[illegible]

Capitation and Medicare

Joseph P. Newhouse

REPORTS
R
728
.6
N43
1986

RAND

The research described in this report was supported by the Health Care Financing Administration, U.S. Department of Health and Human Services, under Grant No. 99-C-98489/9-03.

ISBN: 0-8330-0768-8

The RAND Publication Series: The Report is the principal publication documenting and transmitting RAND's major research findings and final research results. The RAND Note reports other outputs of sponsored research for general distribution. Publications of The RAND Corporation do not necessarily reflect the opinions or policies of the sponsors of RAND research.

Published by The RAND Corporation
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90406-2138

R
728.6
.N43
1986

Sherry A. Terrell, Ph.D.
N1SB/Dres/or/Ord/HCFA
6325 Security Blvd. 2B140M
Baltimore, MD 21207
(301) 597-1416 FTS987-1416

R-3455-HCFA

Capitation and Medicare

Joseph P. Newhouse

October 1986

Prepared for the
Health Care Financing Administration,
U.S. Department of Health and Human Services

RAND

RAND/UCLA Center
for Health Care
Financing Policy
Research

PREFACE

This report was written as part of the Year 3 research agenda of the RAND/UCLA Center for Health Care Financing Policy Research, sponsored by the Health Care Financing Administration (HCFA), U.S. Department of Health and Human Services. HCFA requested that the Center examine the formula used at present by Medicare to pay Health Maintenance Organizations and other alternative delivery systems, and suggest possible improvements. The report, a shorter version of which will appear in the 1986 Annual Supplement to the *Health Care Financing Review*, is the product of that request.

SUMMARY

Increasing numbers of Medicare beneficiaries are receiving medical care from organizations that are paid a fixed monthly amount to treat them, irrespective of the amount of services the enrollees use. The fixed monthly amount is called a capitation, and current Administration policy favors continued growth in capitation arrangements at the expense of traditional fee-for-service medicine.

The growth of capitation has highlighted certain inadequacies in the methods Medicare uses to set the fixed monthly amount, inadequacies that were of little practical import in the past when only a tiny fraction of Medicare beneficiaries had their care paid for in this fashion. Three inadequacies are discussed in this report.

First, the present method ties the capitation to costs in the fee-for-service sector. If the health or medical status of those remaining in the fee-for-service sector differs from those with capitated arrangements, the fixed monthly fee will be either too large or too small. To remedy this problem requires, among other things, adjustments to the fee-for-service costs that take reasonably complete account of health differences between the two groups.

Second, the present method, which is based on fee-for-service costs by county, appears to cause too much spread in rates across counties. This could be remedied by using shrinkage estimators applied to mean costs by county.

Third, the present method for setting a price adjusts for the age, sex, welfare status, and institutional status of capitated enrollees, but this list of variables appears most inadequate; in particular, these variables explain less than 1 percent of the variation in annual per-person expenditure. A good set of adjusters would explain most of the variation in expected (not actual) expenditure at the individual level. (This ignores the separable issue that Medicare does not wish to pay for non-efficacious care.) Expected expenditure has less variation across individuals than actual expenditure, perhaps only a fifth as much, perhaps more. Thus, a good set of adjusters should probably explain at least 20 percent of the variance in actual expenditure. Clearly, the present set of adjusters is a long way from meeting such a standard.

Without better adjusters there is a priori reason to suppose that individuals with high expected expenditure will experience access problems; that is, because they are unprofitable, they will find few providers who will treat them. Evidence of actual access problems is scanty, but some evidence suggests that healthier individuals enroll in capitated systems.

Improving the situation (i.e., finding good adjusters) will not be easy. The two generic types of variables most commonly mentioned as additional adjusters are health status and past utilization. Adding health status variables would require a major research effort. Such an effort would have to address at least three serious problems: (1) Subjective measures of health status may be vulnerable to manipulation (fraud) if they are used as a basis of payment. (2) If treatment changes the value of a health status measure (objective or subjective) and continuing treatment is required (e.g., taking antihypertensive medications), one wants to know the value of the measure before treatment (e.g., the uncontrolled value of blood pressure). But it is generally both dangerous and unethical to try to observe such a measure. (3) Data collection costs are probably nontrivial unless the provider reports the health status measure. In that case, however, there would be an audit problem, because health status measures change through time for reasons related to both natural causes and medical treatment.

Successfully incorporating health status measures as adjusters seems several years off, at best. As a result, a number of individuals have suggested incorporating utilization into the payment formula. The usual measure proposed is past utilization, but current utilization would seem preferable; it is more sensitive to changes in health status, and it avoids the problem that no prior utilization data are available for each new cohort of enrollees. The incentive effects of prior and current utilization are (or can be made) very similar.

Some object to using measures of utilization on the grounds that they give wrong incentives; in particular, they reward a more elaborate style of care. However, pure capitation also has undesirable properties if the amount does not reflect the expected cost of treating the patient. In practice, at least in the short run, there is no ideal system; the problem is to choose the least flawed system. In this situation a blend of fee-for-service and capitation methods has much to recommend it. It is certainly less risky than a pure capitation strategy. Moreover, as methods for adjusting the capitation formula improve, the weight on the capitated amount could increase. The weight on the capitated amount could also be higher in more homogeneous categories of patients (perhaps the younger elderly). How the weights should be determined, however, is a topic left for future research.

ACKNOWLEDGMENTS

I would like to thank my RAND colleagues Stephen Carroll, Shan Cretin, Willard Manning, and Richard Neu, as well as James Lubitz of the Health Care Financing Administration, for helpful comments on a preliminary draft. They, of course, are not responsible for any remaining errors, nor do they necessarily subscribe to the ideas presented here.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS vii

TABLES xi

Section

I. INTRODUCTION 1

II. THE ADJUSTED AVERAGE PER CAPITA COST (AAPCC) AND THE LINK WITH FEE-FOR-SERVICE 2

The Average Community Rate 3

The Method Used to Estimate Costs by County in the AAPCC 5

The Problem of Adjusters 6

Evidence on Biased Selection 13

New Adjusters in the AAPCC Formula 17

III. WHERE TO FROM HERE? 24

Goals and Means 26

REFERENCES 29

TABLES

1. Costs in 1975, by 1974 Use 12
2. Hospital Days of Care per Thousand Among Those Who Chose an HMO and Those Who Chose to Remain in the Fee-for-Service System, Decision Made in October 1976 14
3. Expenditure by Age and Prior Disability Status, Medicare Enrollees 65 and Over, 1980 18

I. INTRODUCTION

Until recently, Medicare has made little use of capitation, the practice by which an organization is paid a fixed annual amount per person to provide necessary medical services. Several recent trends, however, have given impetus to increasing Medicare's reliance on capitation.

First, it has become increasingly accepted that group and staff model health maintenance organizations (HMOs), the premier examples of capitated delivery systems, practice a less expensive style of medicine, with little or no adverse effect on health (Luft, 1981; Manning et al., 1984; Ware et al., 1986). Many attribute the cost reduction to the incentive of capitation.[1]

Second, capitation is seen as an answer to the incentives in the Prospective Payment System (PPS) to "unbundle," that is, to shift services out from under the fixed price the hospital receives for the inpatient stay (Morrisey et al., 1984). A clear example of unbundling is the earlier discharge of patients from the hospital to a skilled nursing facility or to home health care, both of which are reimbursed separately from the fixed payment made to the hospital. If Medicare paid one organization an annual amount and looked to that organization to be fiscally responsible for all services, the unbundling problem should be minimal.[2]

Increased reliance on capitation, however, makes more acute the problems involved in setting the capitated rate or price. This report examines three such problems: (1) the present link between the capitation amount and fee-for-service payments; (2) the likelihood that the present method used to calculate the capitation rate introduces too much spread (variance) across geographic areas in the rates; and (3) the characteristics of enrollees in a capitated system that should be accounted for in computing the rate. We will term these characteristics "adjusters," because they are used to adjust the rate. Most of the report is taken up with the last issue.

---

[1]See, for example, President Nixon's February 18, 1971 message to Congress on HMOs, as quoted in Iglehart (1982). See also Iglehart (1985). Although much commentary focuses on the incentives of capitation, some have pointed out that large multispecialty groups, whether prepaid or fee-for-service, may practice with a less hospital-intensive style (Nobrega et al., 1982; Gaus, Cooper, and Hirschman, 1976).

[2]We say "minimal" rather than "nonexistent", because there might still be some attempt to substitute uncovered services, such as drugs, for covered services.

II. THE ADJUSTED AVERAGE PER CAPITA COST (AAPCC) AND THE LINK WITH FEE-FOR-SERVICE

To set the stage for discussing the three issues, some description may be helpful. The original Medicare Act (Title XVIII of the Social Security Act, enacted in 1965) made very little accommodation to capitated delivery systems such as Health Maintenance Organizations (HMOs). Indeed, it had no arrangements for a prospective capitated rate at all, and simply provided that HMOs could be reimbursed at cost. After the 1971 Nixon white paper endorsing HMOs, Congress enacted the 1972 amendments to the Social Security Act, which provided that Medicare could sign at-risk contracts with HMOs (i.e., could use a prospective capitated rate). The regulations implementing this legislation, however, were not issued until late 1976. They defined the Adjusted Average Per Capita Cost (AAPCC), 95 percent of which equalled the capitation rate. Conceptually, the AAPCC was the estimated cost to Medicare if the HMO enrollees had instead remained in the fee-for-service system. If HMO costs were less than 95 percent of the AAPCC, the HMO could keep half the difference, up to a maximum of 10 percent of the AAPCC.

The AAPCC is calculated by county. It begins with the cost per fee-for-service Medicare enrollee in that county.[1] It then adjusts that cost to the extent the HMO enrollees differ from the fee-for-service population of the county in terms of age, sex, welfare status, and institutional status.[2] The adjustment factors are based on national (fee-for-service) experience.[3] For example, there are five age groups considered by the AAPCC, and the adjustment factor for each age group reflects expenditures by that age group in the national program.

Medicare's ability to contract with HMOs on an at-risk basis did not much enhance its attractiveness to HMOs; apart from demonstration projects, by 1982, when new legislation was enacted, only two HMOs had agreed to be at risk for Medicare beneficiaries.[4]

---

[1]The statement in the text is a slight oversimplification. The AAPCC multiplies the national per capita cost in any year by the ratio of the county per capita cost to the national per capita cost, but the latter ratio is defined over a five-year moving average.

[2]See Kunkel and Powell (1981), Beebe, Lubitz, and Eggers (1985), or the Federal Register (1985) for a technical description of the AAPCC.

[3]The three categories of Medicare beneficiaries—the aged, the disabled, and those with end-stage renal disease—are treated separately in making these calculations.

[4]One unattractive feature was that the calculation of the AAPCC and of HMO costs were both done retrospectively, so HMOs did not know the amount of the payment when they enrolled people.

The AAPCC concept is still employed by Medicare in modified form, and, as the number of those enrolled in capitated systems expands, a difficulty is becoming increasingly evident. By basing the capitation rate on the utilization of those remaining in the fee-for-service system, the AAPCC reimbursement method is vulnerable if capitated systems obtain a nonrepresentative group of enrollees within the AAPCC's age, sex, welfare status, and institutional status categories. In a given year, for example, HMOs might increase their enrollment of low-cost individuals within each category (so-called favorable selection). If this were to happen, the Medicare program would overpay HMOs in two senses: (1) Medicare would pay an average rate (actually, 95 percent of the AAPCC) for a below-average cost group; and (2) when the AAPCC is recomputed next year, the average for those remaining in the fee-for-service system will be higher, and so next year HMOs will receive even more. Thus, total Medicare outlays will rise. Clearly, the converse problem could also arise: Medicare could underpay HMOs if they enroll high-cost individuals within each category. The difficulty arises because the AAPCC links the HMO payment to what Medicare pays for fee-for-service enrollees.

THE AVERAGE COMMUNITY RATE

The Tax Equity and Fiscal Responsibility Act (TEFRA) of 1982 caused the regulations to change again (in 1985) by introducing the concept of the Average Community Rate (ACR), which partially addresses the vulnerability of the AAPCC just described. The ACR is based on rates an HMO or alternative delivery system charges in its private business (for patients under 65 years of age), adjusted for differences between the privately insured population and the Medicare population (Trieger, Galblum, and Riley, 1982). The HMO is still paid 95 percent of the AAPCC, but if that amount exceeds the ACR, the difference must be used to provide Medicare enrollees additional benefits or reduce their cost sharing.[5] If 95 percent of the AAPCC is less than the ACR, however, the HMO need not extend additional benefits. Moreover, HMOs no longer have to share "profit" with the government.

Because the ACR is determined by what HMOs charge in their private business, and not what Medicare pays fee-for-service providers, it is unaffected by HMOs enrolling a nonrepresentative (within age-sex

[5]The HMO could accept reduced payments, but there is clearly no reason to do so.

categories) group of Medicare beneficiaries.[6] Medicare payments are still vulnerable to such nonrepresentative enrollment, however, because it affects the AAPCC and any excess of 95 percent of the AAPCC over the ACR is funneled back to beneficiaries.[7]

Perhaps more important, the promise of the ACR to pay alternative delivery systems on the basis of a market rate rather than an administratively determined rate appears negated in practice because of important implementation problems. The difficulty arises from the adjustments necessary to make the Medicare rate comparable to the HMO rate for the under-65 population.

Adjustments must be made for both differences in benefits between the Medicare program and private business, as well as the higher rates at which the elderly use any given benefit.[8] Consider first the benefit differences. Some arbitrariness can be introduced in costing benefits the HMO provides in its private business that Medicare does not provide. For example, an HMO that provides its under-65 population a prescription drug benefit has an incentive to minimize the "cost" of that benefit because Medicare does not cover outpatient drugs, and to maximize the cost of hospital and physician services that Medicare does cover.[9] The problem becomes even more acute if Medicare covers a benefit that the HMO does not provide in its private business (or provides to few individuals); in that case Medicare does not observe a market rate.

In practice, however, the more serious problem is probably the adjustment for the higher use of the elderly. This adjustment has in fact been based on the observed use of the elderly relative to the non-elderly within any given HMO. Such an adjustment clearly moves away from the notion of a market price for the elderly that the HMO takes as a given. If, for example, the HMO decides to increase its services by 10 percent to the elderly, while keeping constant its services to the non-elderly, its reimbursement would rise 10 percent, up to a limit of 95 percent of the AAPCC. Using factors based on national experience to adjust the HMO's rate would eliminate this problem, but would raise another: Both the aged and the non-aged who are enrolled in any

---

[6]If HMOs enroll sickly individuals, such individuals also are vulnerable.

[7]Regulations preclude HMOs from charging a sham rate on a tiny volume of private business by a requirement that half (or more) of the enrollees not be Medicare or Medicaid beneficiaries.

[8]The elderly use services at more than three times the rate of the non-elderly (Fisher, 1980).

[9]In particular, it will seek accounting rules that allocate as much joint cost as possible away from services such as drugs; for further analysis of this problem, see Danzon (1982).

given HMO may differ from the national population. Thus, use of national experience would raise the same problem that now plagues the AAPCC, namely, how to adjust rates for varying characteristics of those enrolled in the HMO.

THE METHOD USED TO ESTIMATE COSTS BY COUNTY IN THE AAPCC

Before coming to the issue of the adequacy of the adjusters used in the AAPCC, let us briefly examine a separable issue. As noted above, the AAPCC is based on mean fee-for-service costs in the enrollee's county of residence. Some have suggested that the resulting costs seem to have too much spread. Greenlick (1985), for example, notes that the five AAPCCs in the five counties that make up the Portland (Oregon) metropolitan area vary by a factor of two, and it does not seem likely that expected annual costs of caring for Medicare beneficiaries vary that much among contiguous counties.[10]

Part of the problem may be the use of the (sample) county mean to estimate the AAPCC for the county. Consider what turns out to be an analogous problem, that of estimating the end-of-season batting averages of baseball players from their first 50 at bats. Those who follow baseball know that there will be more spread in batting averages after only 50 at bats than at the end of the season. Although it is moderately unusual to have .400 hitters after 50 at bats, it is truly extraordinary to have any .400 hitter at the end of the season. (Such an average has not been achieved for 35 years.)

Recent advances in statistics show that if one wants to predict the batting average of each player at the end of the season, one can do better than simply using his average after 50 at bats. Similarly, if one wants to estimate the true (population) mean expenditure rate for many counties, one can do better than use each county's sample mean.

The better method uses a weighted combination of the individual county mean and the mean of a larger population (the metropolitan mean, for example, or the mean of several rural counties), where the weight on the individual county mean is less, the less the spread among all the county means and the more county means are being considered. Such a procedure "shrinks" all the county means toward the mean of

[10]Part of this variation, perhaps a large part, reflects "style" differences among providers (Wennberg and Gittelsohn, 1973; Chassin et al., 1986). The shrinkage estimator proposed here does not address that issue, but rather the problem that the actual mean costs per county in any year may be "noisy" estimates of expected costs because of the skewness of the expenditure distribution.

the larger group (i.e., it will reduce the spread in the county means), and it can be shown to have lower mean square error for estimating the true mean of each county than using the sample mean for each county.[11] How this method might be implemented in the Medicare program is an issue for another paper.

One may ask why the county is the appropriate geographic area to use in the first place, rather than, say, the mean of a larger area, perhaps a metropolitan area. Essentially, the choice of area size is a tradeoff between the reduction in variance (or noise) one would obtain with a larger area versus the (likely) increased heterogeneity within the area; for example, the mean of a metropolitan area would mask any true differences in the means among the various counties it comprises. Exploring this tradeoff would be a worthy future research project.

THE PROBLEM OF ADJUSTERS

What variables should be included in the AAPCC? More generally, what variables should be used to adjust for differential risk of treatment among people who are enrolled in a capitated health plan?[12] As noted above, the present AAPCC uses age, sex, welfare status, and institutional status, but there is near universal agreement that this list is inadequate because it explains only 0.6 percent of the variance in Medicare expenditure (Lubitz, Beebe, and Riley, 1985). Although 0.6 percent is on its face too low a number, it is less clear what amount of variance one should seek to explain.

In principle, one wants to estimate the expected expenditure for each individual; actual expenditure in any year will differ depending on how sick the person is that year.[13] Thus, one wants to know how much

---

[11]For a nontechnical account, see Efron and Morris (1977), who use the batting average example. A very similar problem to estimating the AAPCC for each county is to estimate the true rate of false alarms from any given fire alarm box. The solution to this problem is given in Carter and Rolph (1974). Methods proposed by Efron and Morris (1972) are also relevant for this problem.

[12]The issue of accounting for heterogeneity arises irrespective of how a base price or average price is set. The issue of setting a base price is not taken up here, except for the following comment. The AAPCC uses fee-for-service system prices to set a base price. The main alternatives involve the alternative delivery system's quoting a price. Either the beneficiary could pay this price, with partial or full reimbursement coming from Medicare (i.e., a voucher system), or Medicare could contract with certain delivery systems (i.e., bidding). In both cases, however, the issue would arise of how the price (or value of the voucher) will vary with the characteristics of the beneficiary. That is the issue considered here.

[13]We do not take up the normative issue that Medicare may not wish to pay for some expenditure (e.g., iatrogenic or nonefficacious care). To keep the discussion focused on the problem of heterogeneity among individuals, the actual expenditure by each individ-

expected expenditure varies across people. The remainder of this section is mildly technical, the intent being to suggest methods for estimating variation in expected expenditure when one observes variation in actual expenditure.

In order to discuss variation in expected expenditure, it will be necessary to introduce some notation. The discussion will be kept as accessible as possible, however, by organizing it around a model of expenditure that has been widely used in the literature about this problem (e.g., Newhouse, 1982; McCall and Wai, 1983; and Welch, 1985a, 1985b). In this model, called the components of variance or error components model, medical expenditures are predicted by the following type of equation, which includes both adjusters and two components of an error term:

$$\text{Expenditure of person } i \text{ in year } t = \beta(0) + \beta(1)(\text{Age}) \quad (1)$$

$$+ \beta(2)\text{Gender} + \beta(3)\text{Welfare Status} + \beta(4)\text{Institutional Status}$$

$$+ \mu(\text{i}) + \epsilon(\text{it}).$$

The β's are constants that multiply the various adjusters. For example, the adjuster for gender equals 1 if the person is female and equals 0 if the person is male; hence, average expenditure by females differs from males by an amount $\beta(2)$ (holding constant age, welfare, and institutional status).

This illustration has included the four variables that are in the AAPCC as adjusters,[14] but one could add any measurable variable as an adjuster (or could drop one or more of these four). As variables are added or dropped, the β's will generally change. Inevitably, however, not all relevant characteristics will be used as adjusters. Some may be exceedingly costly to measure; others may have such a small role to play that it is not worth including them.

The characteristics that are not included as adjusters are incorporated in the $\mu(\text{i})$ and $\epsilon(\text{it})$. The effects of variables included in the $\mu(\text{i})$ term vary across people, but for a given person they do not vary with time; we relax this somewhat restrictive assumption later. For the average person $\mu(\text{i})$ is zero, whereas for a hypochondriac $\mu(\text{i})$ is most likely positive, and for a stoic most likely negative.

---

ual will be treated here as proportional to (more specifically, a linear function of) the care Medicare wishes to pay for.

[14]The actual AAPCC comes from a so-called fully interacted specification, rather than the main-effects specification described here, but that difference is unimportant for present purposes. Age is specified as having five intervals in the actual AAPCC.

By contrast with the $\mu(i)$ term, the effects of variables included in the $\epsilon(it)$ term do vary with time; they represent the consequences of events specific to individual i in year t that cause him or her to have high or low expenditure in that year (high or low relative to his or her mean). If it is an "average" year for the person, $\epsilon(it)$ will be zero. But if the individual falls and breaks a hip in that year, $\epsilon(it)$ would probably be quite positive.

It is important to distinguish $\mu(i)$ and $\epsilon(it)$ because $\mu(i)$ enters an individual's expected expenditure, whereas $\epsilon(it)$ has an expected value of zero for each individual. Thus, an alternative delivery system could plausibly estimate $\mu(i)$ in advance, whereas it is more difficult to estimate $\epsilon(it)$ in advance. In other words, although Medicare does not use, as adjusters, variables whose effects are summarized by the $\mu(i)$, the alternative delivery system can observe, or at least estimate, their effects once the individual is enrolled. For example, the organization can observe race, and it will learn if the enrollee is a hypochondriac. The latter is an example of a type of variable that may be impossible for Medicare to measure and hence makes adjustment a quite difficult task.

Nonetheless, adjustment is important because enrollees with negative $\mu(i)$ are desirable because they will be "profitable." By contrast, individuals with positive $\mu(i)$ will be unprofitable. If those individuals are known in advance, an HMO or alternative delivery system has an incentive to persuade them to seek care from another organization, just as a person playing the game of Hearts tries to send the Queen of Spades elsewhere. The term "selection behavior" will be used here to refer to behavior that a delivery system may engage in to attract individuals with negative $\mu(i)$ and repel those with positive $\mu(i)$.[15] If some systems engage in selection behavior, certain individuals will find they are unwelcome in any capitated system, because those systems that do not engage in such behavior will find themselves losing money. As this happens, they too will be forced to discriminate among risks, reduce services to the sickly patient (but not to the not-so-sick patients), or go out of business.

Medicare (HCFA) in fact attempts to prohibit such behavior through regulation, but such regulations appear unenforceable. If a physician wants to discourage a patient from seeing him or her, there are many subtle ways to do so (e.g., appearing unsure about the

[15]This may sound like unlikely behavior, but under capitation there can be a strong monetary incentive to engage in selection, whereas under traditional fee-for-service there has not been such a financial incentive to do so. Individuals themselves may engage in selection behavior if their premiums reflect the cost of the average person enrolled in (or insured by) the organization.

patient's diagnosis, being curt, keeping the patient waiting). If there is a residual fee-for-service system, such individuals would presumably remain in that system. In such a case (under the usual assumptions), those whose expected costs are less than the capitated rate will enroll in alternative delivery systems, which make extra profits on all but the marginal case; these profits may be "competed away" in amenities (Held and Pauly, 1983). In any event, expenditures by the government rise.

How Serious a Problem is Selection Behavior?

The issue of selection behavior is not new; the literature has devoted considerable attention to it. The most comprehensive quantitative discussions to date have been those of Welch (1985a) and Lubitz, Beebe, and Riley (1985). Although Lubitz, Beebe, and Riley present data suggesting that selection may well be a problem (some of their data appear below), Welch's general tenor is optimistic; he, in effect, argues that the quantitative magnitude of the problem is small. The present author is less sanguine than Welch, but to explain why, we must examine his two arguments:

1. Because the variance of $\mu(i)$ is a relatively small portion of the variance of the total error term $\mu(i) + \epsilon(it)$, the problem is not as daunting as it might first appear. Specifically, the variance of $\epsilon(it)$ among the under-65 is about four times the variance of $\mu(i)$ (Newhouse, 1982).[16] Because Welch believes one cannot generally predict $\epsilon(it)$, the maximum proportion of variance one could explain by a set of adjusters is around 0.2.[17] Although adjusters that explained 16 percent of the variance (i.e., R-squared is 0.16) may seem inadequate, in fact one would have explained 80 percent of the predictable variance (i.e, the variance that might be used by an alternative delivery system to select preferred risks).

---

[16]McCall and Wai (1983) estimate that for total expenditure the variance of $\epsilon(it)$ is seven times that of the variance of μ. The figure, of course, depends on what variables are included in the observable variable list. Nonetheless, the values are reasonably consistent. The average 0.2 figure may differ by subgroup.

[17]That is, the maximum R-squared one could attain in explaining annual expenditure is about 0.2. The argument follows from the facts that (1) the current AAPCC adjusters explain less than 1 percent of the variance, and (2) that other adjusters one might add to the formula could only explain variation that is now in the $\mu(i)$ (on the assumption that one cannot predict the $\epsilon(it)$). Welch notes that including year-specific data such as maternity that are not under the control of the HMO would be a useful addition to the AAPCC. With such variables, one could expect to do better than explain 20 percent of the variance because one is predicting variation in $\epsilon(it)$. This argument is correct; we return to it below.

2. Welch's second argument introduces a modification in Eq. (1) because the data suggest that the value of $\mu(i)$ changes over time; in particular the data suggest that $\mu(i)$ is likely to be closer to zero next year than it is this year. Hence, the value to an organization of selecting individuals on the basis of any particular value of $\mu(i)$ steadily decays; in effect there is less to be gained by trying either to select good risks or discriminate against bad risks because good risks may become bad risks or bad risks may become less bad.

Welch assumes that $\mu(i)$ varies with time in the following way:[18]

$$\mu(i,t) = \rho * \mu(i,t-1) + \epsilon(i,t), \tag{2}$$

where ρ is a constant between zero and one.

Based on data in Beebe, Lubitz, and Eggers (1985), Welch estimates that ρ is around 0.89, implying that any given value of $\mu(it)$ can be expected to decay to half its present value in a 5-to-6-year period.

Although Welch's arguments are correct, they give little basis for optimism. There is first the obvious problem that the current adjusters are a long way from predicting even 20 percent of the variance; specifically, they explain only 3 percent of the variance that Welch considers predictable.

A less obvious problem occurs in the assumption that the variance in the $\epsilon(it)$ term is unpredictable and can be ignored. If a delivery system can predict $\epsilon(it)$, the scope for selection of patients obviously becomes much broader.[19]

Welch does note that in the case of a pregnant woman, a high value of $\epsilon(it)$ is predictable. That is, the delivery system can expect that in the future a pregnant woman will probably incur above-average medical costs. He notes that, as a result, pregnancy probably should be added to any AAPCC used for Medicaid beneficiaries. In general, however, he seems to regard pregnancy as an isolated example and the issue of predicting $\epsilon(it)$ of little relevance to the Medicare program.

Unfortunately, pregnancy is not such an isolated example. The outstanding analogy of relevance to the Medicare population is that of enrollees at a high risk of death. Although only 5 percent of all Medicare enrollees die in any year, they account for 28 percent of all Medicare expenditures in the year (Lubitz and Prihoda, 1984). Of the 5 percent who die, many may give no early warning, but a considerable

[18]The model in Eq. (1) is obviously a special case of this model in which ρ equals 1.0.

[19]Technically, if $\epsilon(it)$ is predictable, the fact that the variance of the $\mu(i)$ is only about a fifth of the total error variance is of less relevance; because $\epsilon(it)$ is observable in time $t + 1$, it is of less relevance that the $\mu(i)$ follow an autoregressive process.

fraction may be known to the delivery system to be at high risk of death.[20] Moreover, the deaths with considerable warning may be those with disproportionately high expenditure. Someone who dies suddenly and without warning from a heart attack does not utilize much care.

Equally important for this discussion, peak spending tends to occur in the last few weeks of life. Of all expenses in the last year of life, 30 percent are spent in the last 30 days and another 16 percent in the 60th to the 31st day before death (Lubitz and Prihoda, 1984). In other words, nearly half the expenditures on those who die occur in the two months preceding death. Of course, not all patients at high risk of death die. However, those who do not are almost as expensive as those who do (Detsky, Stricker, Mulley, et al., 1981). Thus, like a pregnant woman, an individual who is likely to die shortly will usually have future expenditures that are above present expenditures.

Not only is part of the variation in ϵ(it) predictable ex ante, but by definition it predicts μ in the following accounting period.[21] For example, patients may experience an acute condition or a flareup of a chronic condition that will require treatment in the next accounting period, in which case expenditure in the next twelve months may be predictably well above average.

Data from Anderson and Gruenberg (cited in Lubitz, Beebe, and Riley, 1985) give some feel for the magnitude of this problem. They categorized medical expenditure in 1975 among women aged 70–74 according to their use in 1974, and presented the data shown in Table 1.

The last two rows of Table 1 exemplify large positive values of ϵ(it) that would be observable by the delivery system. In other words, at the end of 1974 the delivery system would know that individuals have had a hospitalization of this type and hence will have an expected value for μ(i,1975) that will be quite positive. Thus, the delivery system will have an incentive to purge itself of the enrollee.

In sum, the potentially large predictable components in μ(it) and ϵ(it) suggest that the task of refining the AAPCC is important and will not be simple.

---

[20]The author is not aware of any research that has established such a fraction for the general population.

[21]The term "accounting period" is used here rather than year to emphasize that an adverse health event may occur near the end of an accounting period (i.e., close to open enrollment), in which case the rate of spending in the next accounting period may be well above the average rate in the present period.

Table 1

COSTS IN 1975, BY 1974 USE

1974 Usage Group	1975 Costs Relative to the Average Medicare Enrollee
No use	0.34
Part B use only	0.82
Nonchronic hospitalization	1.16
Hospitalization for chronic condition	1.92
Repeated hospitalizations for cardiac, or musculoskeletal conditions or cancer	3.45

A Contrary View

On the basis of the data presented above, the case that the current version of the AAPCC needs modification seems compelling and is, in fact, widely accepted. Nonetheless, a contrary view holds that the problem of selection is exaggerated (Pauly, 1984, 1985). Pauly believes that the empirical evidence on the actual extent of biased selection is weak; that evidence is discussed in the next section. Pauly also argues that the theoretical case for inefficiency as a result of selection is based on a model described by Rothschild and Stiglitz (1976), the inefficiency therein being that good health risks cannot obtain as much insurance as they wish. Pauly correctly notes that no one has presented evidence that establishes such an inefficiency, and he draws the overall conclusion that "as long as one avoids things like community rating and too easy switches across policies, adverse selection need not be a difficulty."

Pauly's dismissal of adverse selection follows logically from his premise, but the premise is of doubtful validity. "Avoiding community rating" is taken here to mean avoiding paying the same amount for a group with different expected costs; thus, either the capitated rate is set individually, or nearly homogeneous risk groups are defined. The evidence presented above has suggested that we are a long way from that.[22] Indeed, the AAPCC is very close to community rating (since the

[22]If the alternative delivery states a price for each individual and attempts to include $\mu(i)$ and the predictable portions of $\epsilon(it)$ in that price, then selection should not be an

current adjusters explain less than 1 percent of the variance in expenditure).

Although community rating is the principal issue, the avoidance of "too easy switches" in the Medicare program is also an issue; enrollees can at present change plans every month. This regulation on plan changing probably should be changed to once a year, the general standard in the private sector. In sum, because the Medicare program can scarcely be said to have avoided community rating (to date), and because plan switches by Medicare beneficiaries seem easy, we have no reason on a priori grounds to be optimistic about selection under current arrangements.

Some object to individual rating on the grounds that it fails to insure against bad luck at birth; that is, sickly individuals would face higher premiums because they have higher expected costs. If desired, however, more sickly individuals could be given higher vouchers; in any event, this issue of equity is not the issue being addressed here. Rather, the focus is on the extent to which individual rating is feasible—what adjusters can be used—and the consequences if the adjusters actually used are not very good.

EVIDENCE ON BIASED SELECTION

The argument to this point has been that we have some indirect or theoretical reasons to expect that adverse selection could be a serious problem. Such an argument certainly is not as compelling as direct evidence; but what evidence is there that adverse selection is an important empirical problem? Unfortunately, there is rather little direct evidence one way or the other. Much of it comes from comparing those who either enroll or disenroll from HMOs with those who remain in whatever system they have chosen, and these two generic types of studies yield conflicting evidence.

---

issue. This appears to be what Pauly has in mind. However, the author knows of no actual example of attempts to rate each person individually; invariably, prices are quoted for classes of enrollees (viz., the AAPCC) (see Newhouse, 1984). In most cases studied in the literature (the evidence Pauly believes is weak), there is little or no attempt to adjust at the individual level. Studies of other insurance markets, such as the automobile insurance market, could suggest how the performance of the market improves with the ability of adjusters to explain variation at the individual level (e.g., automobile insurers use age, sex, type of automobile, place of residence, miles driven, and, most important, accident experience, in setting premiums).

Evidence on Enrollment

Several studies have compared use-rates of HMO enrollees and non-enrollees at a time prior to HMO enrollment, when both groups were in the fee-for-service system (Eggers, 1980; Eggers and Prihoda, 1982; Luft, 1981; Jackson-Beeck and Kleinman, 1983; Buchanan and Cretin, 1986). These studies tend to show that when HMO enrollees and non-HMO enrollees were both enrolled in the fee-for-service system, the HMO enrollees used less care.[23] For example, Eggers (1980) compared the hospitalization rates of Medicare beneficiaries who enrolled in an HMO with those who remained in the fee-for-service system at a time just prior to their choice, when both groups were in the fee-for-service system; the factor-of-two difference between the two groups shown in Table 2 suggests strong selection.

Despite Welch's conclusion that the observed differences between people who choose to enroll in HMOs and those who do not will subsequently narrow, the magnitude of the differences in studies of enrollees seems large enough to be worried about the spread of capitation without modifications to the current AAPCC.[24]

Table 2

HOSPITAL DAYS OF CARE PER THOUSAND AMONG THOSE WHO CHOSE AN HMO AND THOSE WHO CHOSE TO REMAIN IN THE FEE-FOR-SERVICE SYSTEM, DECISION MADE IN OCTOBER 1976

Year	Persons Choosing HMO	Person Choosing Fee-for-Service
1974	1152	incomplete
1975	849	1761
1976	731	1929

[23]Buchanan and Cretin (1986) note that in the year immediately preceding HMO enrollment, utilization was even lower than in the years before that for HMO enrollees, suggesting that they were postponing utilization until they were enrolled in the HMO.

[24]For example, an infinite series whose value in each subsequent period is 0.89 times the value in the present period (0.89 is Welch's estimate of ρ) at a 10 percent discount rate is still worth about half as much as an infinite series that does not decay (i.e., has a ρ of 1.0).

Evidence on Disenrollment

If selection is a problem, one might expect to see that disenrollees from HMOs use more services than those who remain in HMOs. A number of studies of disenrollees have been conducted (Wollstadt, Shapiro, and Bice, 1978; Wersinger and Sorenson, 1982; Mechanic, Weiss, and Cleary, 1983; Hennelly and Boxerman, 1983; Griffith, Baloff, and Spitznagel, 1984; Lewis, 1984; Buchanan and Cretin, 1986). In contrast to the studies of enrollees, most studies of disenrollees show little, if any, evidence of adverse selection. (Buchanan and Cretin (1986) is an exception.) Can we conclude from these studies that adverse selection is a problem only in theory and that we can proceed toward capitation without much concern for additional adjusters?[25] The answer is most likely no.

One should distinguish four reasons for disenrollment: (1) mandatory disenrollment—the individual moves out of the HMO service area, or changes employers, and the HMO is not offered by the new employer; (2) disenrollment because of a change in the relative prices of the HMO and fee-for-service options (or perhaps some change in the nonprice amenities of the two systems of care, such as the HMO's closing a clinic); (3) disenrollment because the person made a "mistake" in enrolling; and, finally, (4) selection, meaning HMO-induced disenrollment of those with expenditures expected to be above average.

One should expect that large samples would be needed to detect adverse selection because the first three causes of disenrollment could tend to obscure it. The reason may become clearer by considering each category in turn.

Mandatory disenrollment can be regarded as outside the control of the HMO. Hence, for purposes of detecting selection (signal), it acts like noise. Some studies distinguish mandatory selection, but not all do.

Those who disenroll because of a change in relative prices are those who are nearly indifferent between enrollment in fee-for-service and the HMO (e.g., those who live quite far from an HMO clinic or who are just barely able to "tolerate" an organized system of care for the monetary savings). The nearly indifferent group may be only weakly related to the group that the HMO would like to see disenroll; hence, it too may act mostly like noise for purposes of detecting selection.

The third category of disenrollees, those who made a mistake, may well be the largest group of disenrollees in any given period (especially

[25]All of these studies were of non-Medicare populations and the AAPCC was not involved. These studies are relevant, nonetheless, because few, if any, adjusters were included in the premium.

if relative premiums have not changed). It is probably impossible to be very certain, without actually joining, just how much one will like care at the HMO,[26] consequently, some people will join, decide they made a mistake, and withdraw. This is especially true of new HMOs; for well-established HMOs, one may be able to form an idea of how well one will like the HMO by speaking with current members.[27]

Detecting HMO-induced disenrollment is the more difficult because of the skewness of the expenditure distribution. The HMO need only induce a few people to leave in order to have a substantial effect on profitability (1 percent of the population accounts for about a quarter of the expenditure on medical care), but the same skewness makes it difficult to detect a difference in spending between those who leave and those who stay; differences in medical expenditures between two groups tend to have large standard errors unless samples are on the order of several thousand individuals. Only quite large HMOs will have that many disenrollees in a year. The studies in the literature mostly have samples of disenrollees in the hundreds.[28] The study with the largest sample (Buchanan and Cretin, 1986) also was the study that found evidence of adverse selection among those who disenroll.

In sum, the studies of disenrollment may simply have failed to detect selection behavior. We cannot be very confident from their negative results that selection behavior is an unimportant problem.

Other Evidence

The RAND Health Insurance Experiment randomized individuals who were receiving care in the fee-for-service system into an HMO. When analysts compared their use-rates with those already enrolled in the HMO, they found virtually no difference, implying no selection (Manning et al., 1984). For present purposes, however, this finding

---

[26]In the jargon of economics, HMOs are an experience good.

[27]For this reason it is not surprising that Wollstadt, Shapiro, and Bice (1978) found that the first cohort of enrollees in a new HMO exhibited the highest disenrollment rates.

[28]Griffith, Baloff, and Spitznagel (1984) have a sample of 2263 voluntary disenrollees from one HMO. They show significantly fewer physician visits among disenrollees, as does Lewis (1984). They interpret this as evidence against selection, but the case is weak, in part because out-of-plan use is not included, in part because the variable measured is visits, not dollars, but primarily because the result could be true if the second and third groups described above have lower true mean rates of visits. Although this appears to leave the HMO with a poorer than average risk mix, so that it is in fact underpaid, any enrollees with above-average expected expense will still find that every organization will have an incentive to purge them. For such enrollees, access could prove to be quite poor. Hence, even if the Griffith, Baloff, and Spitznagel finding is true of other HMOs, there would still be a problem.

must be regarded as merely one observation and is not strong evidence that there would be no selection in a world of competing delivery systems.[29] A related point is that virtually all existing evidence comes from studies of communities with one HMO and a fee-for-service system. As competition among capitated systems increases, selection behavior may well increase.

NEW ADJUSTERS IN THE AAPCC FORMULA

Despite the agreement that the current AAPCC is inadequate, there is considerable disagreement on what should be done about it. This section sketches some issues in this debate; it begins by listing some desirable characteristics of adjusters:

1. Perhaps the most important characteristic is that the set of adjusters should result in reasonably homogeneous categories with respect to expected expenditure, just as the Diagnosis Related Group (DRG) system is intended to be a reasonably homogeneous classification system for hospital cases. This is the criterion that the discussion to this point has emphasized. A necessary condition is that an adjuster predict medical expenditure.
2. An adjuster should not have excessive collection costs.
3. An adjuster should be reasonably cheap to audit or verify, so that it is inexpensive to keep fraud to a minimal level.
4. An adjuster should not induce changes in behavior by patients in order that they be classified in a category with higher reimbursement rates.

Disability Status

Lubitz, Beebe, and Riley (1985) have shown that disability status prior to age 65 predicts medical expenditure. Table 3, taken from their paper, shows that those disabled before age 65 spend 56 percent more after age 65 than those not disabled (1704/1091 = 1.56).

Moreover, Lubitz, Beebe, and Riley (1985) present some evidence of selection on prior disability status. In two of three demonstration HMOs, the percentage of formerly disabled enrollees was 30 and 38 percent less than in a comparison fee-for-service group; in the third HMO the percentages were nearly equal.

---

[29]To conclude from this finding that selection will not occur would be like observing community rating at Blue Cross in the 1930s and concluding that experience rating will not occur.

Table 3

EXPENDITURE BY AGE AND PRIOR DISABILITY STATUS, MEDICARE ENROLLEES 65 AND OVER, 1980

Age	Formerly Disabled	Not Formerly Disabled	All Persons
Total	$1704	$1091	$1127
65-69	1635	748	846
70-74	1761	958	1008
75+	2014	1433	1440

Disability status seems like an almost ideal adjuster. Because it exists in administrative records, the cost of collecting it is small; so too is the cost of auditing it. If it is included, it seems unlikely to affect behavior very much in an undesired way; i.e., it seems unlikely that there is much additional incentive to become disabled before age 65 if the Medicare formula recognizes disability.[30]

The only difficulty with including disability prior to age 65 is a short-run practical problem; the current formula includes institutional status, but there are no data on institutional status in routine administrative data and no data on disability status in the survey data that were used to set the current adjustment factors.

Even if this difficulty is overcome, adding disability status to the formula is likely to increase explained variance by only a modest amount. Lubitz, Beebe, and Riley (1985) do not give data on how much variance disability status explains; however, only about 3 million disabled under 65 were eligible for Medicaid in 1983 (*Health Care Financing Review*, 1984). Hence, the vast majority of aged do not become disabled before age 65, and of course disability status cannot explain any variance within that group.[31] As a result, although there is a good case for adding disability status as an adjuster, it almost certainly cannot come close to "solving" the problem.

[30]If this were a problem, one might make the requirement that one had to be disabled at an earlier age than 65, for example, 60.

[31]Any variance within the disabled group will also remain unexplained.

Health Status and Functional Status

The case for adding health status variables seems straightforward on the surface. Health status measures are known to be important predictors of variance (Manning, Newhouse, and Ware, 1982). For example, individuals with chronic health problems such as cancer or congestive heart failure will tend to have greater values of $\mu(i)$ than individuals without such problems (assuming the present AAPCC adjusters). For that reason, McClure (1984) and Thomas et al. (1983) have strongly argued for research into the means by which health status measures could be added to the formula.

Unfortunately, there are important practical problems in adding health status measures (Lubitz, Beebe, and Riley (1985)). Not surprisingly, these problems are conceptually similar to those associated with the Prospective Payment System (PPS), which can be seen as an effort to incorporate clinical or health status adjusters into the method for paying hospitals. The problems may be briefly summarized as follows:

1. Patients with the same condition might warrant more or less intensive treatment depending on the severity of that condition. For example, a physician might choose not to treat someone with a diastolic blood pressure of 95 mm Hg, but treat immediately someone with a diastolic blood pressure of 125 mm Hg. Yet if there is simply a fixed additional amount for those with hypertension, the payment system is not recognizing any costs associated with treating the more elevated levels of blood pressure (e.g., more frequent followup or other increased intensity of treatment). This problem is analogous to the so-called "severity" problem (within-group variance) in the PPS.

   One solution in principle to the severity problem is to vary the payment with the level of a physiologic measure (e.g., pay more for the person with a diastolic blood pressure of 125 mm Hg). However, treatment may change the level of the physiologic measure. If one pays \$X for someone with a natural diastolic blood pressure of 95 mm Hg, but \$X + D for someone with a reading of 125 mm Hg, one must decide how much is paid if medication lowers blood pressure to 95 mm Hg. Such an individual clearly costs more to treat than an individual with a naturally occurring blood pressure of 95 mm Hg. At first glance it may seem desirable to pay on the basis of the uncontrolled value (in the example, 125 mm Hg), but if the

person's blood pressure is in fact controlled, the uncontrolled value would not be observed—and it would be dangerous and unethical to try to observe it.

A similar problem arises if, as some suggest, payment is based on maintenance or improvement of health status. On average, the health of elderly people can be expected to deteriorate; an organization paid on the basis of maintaining or improving health status faces a Sisyphean task. Conceptually, one would like to pay on the basis of deviations from expected prognosis, but expected prognosis is not observed.

2. Just as it is possible to engage in DRG creep (Carter and Ginsburg, 1985), it appears possible to "game" some health status measures used to adjust capitation amounts, particularly self-reported health status measures. For example, if Medicare pays more for patients who report their health as poor instead of fair, there is an incentive for the patient and the provider to collude against the government to report the patient's condition as poor.[32]
3. Just as co-morbid conditions complicate setting prices for the PPS, they complicate setting prices in a capitated system. An individual who is healthy except for hypertension may be relatively easy to rate. An individual with hypertension, diabetes, kidney failure, and congestive heart failure may require more treatment than merely the adjustments implied by the four diagnoses separately.[33] Just as was done with the PPS, one can simplify by equating all conditions and simply making an adjustment for the presence of any co-morbid condition, but some conditions may be "more equal" than others.
4. Just as implementing DRGs added data-collection expense, some point out that collecting health status data would add expense. Indeed it would, but McClure (1984) argues implicitly that the collection cost issue has been overstated. He notes that delivery systems would need to collect measures of health status for its own purposes, and he would place the onus on the delivery system to notify HCFA of the enrollee's health status. I agree with McClure's judgment that in this case collection costs should be relatively cheap, but he does

[32]Presumably HCFA could prevent the provider's overtly encouraging the patient to give fraudulent answers. Patients may discover, however, that at least some providers give more attention and courtesy to patients for whom they receive more reimbursement. This may well encourage some patients to fudge their answers.

[33]Technically, there may be a very large number of interactions to estimate.

not address the issue of auditing. Auditing may not be a straightforward problem, particularly if the health status reported to HCFA by the delivery system subsequently changes, either for natural reasons or because of treatment. Hence, collection costs do appear to be an issue.

5. Just as the weights assigned to DRGs must be updated to account for new technology, so must capitated payments. The prices paid for new technical procedures, however, are a problem in the present fee-for-service system as well, but perhaps less of a problem because one need not project annual utilization.

These arguments should not impede a research effort on adding health status measures, but they do suggest that a small-scale effort is unlikely to meet much success (see also Lubitz, 1985). Most research to date has focused on the use of functional status as an adjuster (Lubitz, Beebe, and Riley, 1985; Thomas and Lichtenstein, 1986). Functional status describes limitations on mobility and, like disability status, it predicts Medicare expenditure (controlling for the four variables now in the AAPCC). Also like disability status, functional status is clearly observable by the delivery system. Both these considerations argue for its inclusion in the payment formula, but unlike disability status, data on functional limitations do not now exist in administrative records, so special collection efforts would have to be instituted. Moreover, like self-reported health status, data on functional status may be vulnerable to manipulation (the analog of DRG creep).

The Individual at High Risk of Death

Research into inclusion of health status measures as adjusters might begin by finding measures that identify patients at high risk of death. Such patients will often be known to the delivery system to have higher expected expenditure. An organization that chose to engage in selection might well attempt to send such patients elsewhere at a time of open enrollment.

Some argue that an HMO would not want a reputation for "dumping" its patients and therefore would not do so, but if it engaged in this behavior only occasionally, its actions might well not be detected and it would not acquire such a reputation.[34] Nonetheless, if the difference

[34]Anderson and Gertman (1983), cited in Lubitz, Beebe, and Riley (1985), have proposed using mortality rates among enrollees in each system to detect selection of this sort. Some might also propose that mortality rates could be used as an adjuster, because they reflect health status. This has both an incentive and a statistical problem. The

between the expected cost for a patient at high risk of death and reimbursement is great enough, the profits to be made by dumping even a few patients could be substantial. Consequently, the possibility of dumping at least some patients at high risk of death needs to be taken seriously.

The implication is that there should be a higher capitation rate for such patients, but the practical problems of defining such an adjustment remain formidable. Indeed, there is considerable discretion in treating some patients at high risk of death (Garber, Fuchs, and Silverman, 1984), so the magnitude of the adjustment is also a difficult question.

Prior Utilization as an Adjuster

Several analysts (Anderson and Knickman, 1984a, 1984b; Lubitz, Beebe, and Riley, 1985) have investigated adding measures of past utilization to the AAPCC. They demonstrate that past utilization predicts future expenditure; moreover, it is observable both by the HMO and by an agency such as HCFA, and the data are cheap to collect. A different group of analysts (Thomas et al., 1983; McClure, 1984) believes past utilization is inappropriate to include as an adjuster because of its incentive effects. McClure, for example, notes that inclusion of prior utilization will reward a patient of a more "elaborate" provider (or might reward the provider himself or herself) and penalize the patient of a more conservative provider.

The response to this objection on the part of those proposing past utilization as an adjuster is not entirely satisfactory. They note that the adjustment would not be so large that a delivery system could

---

incentive problem is that HMOs with poor results on mortality are implicitly rewarded, while those with good results may be penalized (i.e., they look as if they are dumping). Statistically, an organization that was dumping some patients at high risk of death may stand a good chance of remaining undetected. Suppose there are two HMOs, each with 10,000 Medicare beneficiaries. Suppose 500 (5 percent) of them die annually in both organizations. Now suppose one HMO tries to persuade some patients at high risk of death to move elsewhere. Suppose, to keep the exposition simple, all the patients at high risk of death will die with certainty. If those known to be at high risk of death are 20 percent of all deaths in the Medicare population (100 persons in the example) and one organization sheds a third of those individuals, it loses 33 patients, and its true death rate falls to 4.67 percent. Using a 95 percent chance of Type I error, the chances are only about 1 in 3 in any given year that one will detect a difference between the two organizations with these size samples. (This assumes that the other organization does not gain these individuals, who are dispersed among providers.) While chances would be higher with a larger sample, it is optimistically assumed here that the true mortality rates in the two HMOs are known to be the same (or that one can perfectly adjust for any difference). In practice there would likely be a considerable debate about whether true mortality rates in the two populations were equal (just as teaching hospitals argue for higher PPS payments on the grounds that they treat more "severely" ill patients).

profit by adding more services, by which they mean that it would not pay to take a healthy person and, say, hospitalize him. That argument, however, ignores the fact that decisions are made at the margin. While it will not pay to hospitalize a healthy person, a not-very-sick person may be kept out of the hospital if there is no adjustment for prior utilization, but may be hospitalized if there is such an adjustment. If prior utilization is included in the formula, more elaborate care of a given patient will in general be rewarded; McClure's argument is correct. His conclusion that utilization should be excluded from the payment formula, however, does not necessarily follow.

III. WHERE TO FROM HERE?

McClure (1984) and Lubitz (1985) make a compelling case for a greater allocation of research resources into methods to improve capitation. McClure notes the great disparity between the resources that have gone into developing the Prospective Payment System (PPS) and those that have gone into developing a method for adjusting capitation rates, and argues that at least as sustained an effort as went into the development of DRGs should go into the development of risk-adjusted capitation methods.

The present author agrees, but would caution against taking resources from the effort to refine the PPS to develop capitation methods. Well-developed capitation schemes are at best several years in the future, and in the meantime there are several known technical problems with the PPS now in use. Given the size of the Medicare program and current research budget levels, it is important that efforts to resolve or ameliorate those problems not be diminished.

Part of the greater research effort into capitation ought to include some demonstrations of new capitation schemes (just as the New Jersey experience with DRGs could be thought of as a demonstration of the PPS). The use of functional health status and other measures of health status can be tried out in a demonstration. The introduction of disability status as an adjuster may be straightforward enough as not to warrant a demonstration.

In addition to the research effort, however, consideration should be given to a blend of capitation with fee-for-service. For example, a delivery system might receive a quarter of the current—or better yet, a refined—AAPCC for each person it enrolled, while three-quarters of its payment might be based on fee-for-service. To do so would recognize the imperfect nature of all adjustment methods now available, and although research should improve matters, it may well not produce a workable "solution."[1] Consequently, we ought to give some thought to the situation in which the available adjusters are not fully satisfactory. In such a situation, a blend has much to recommend it; the weight on the capitated portion might increase as adjusters became better.

Moreover, for the period during which we do not have a completely satisfactory set of adjusters, a blend seems to be an improvement over

[1]Analysts at HCFA, for example, believe that "no AAPCC will ever be perfect." (Beebe, Lubitz, and Eggers (1985)). The hypochondriac example, given above, illustrates the problem. Technically, we are in a second-best world.

prior utilization as one of several adjusters.[2] The latter is similar to a blend of capitation and prior use with a weight on prior use given by a regression coefficient (as in Eq. (1)). Use of a blend with current use recognizes changes in health status as they occur rather than with a delay. Moreover, it avoids the problem that no past utilization data are available for a new cohort of enrollees.

On the other hand, those who object to the incentive effects of prior utilization will probably object even more strongly to the use of current utilization as the basis of payment because of its presumed stronger incentive effects. The differences in incentive effects between using prior and current utilization, however, turn out to be more apparent than real. For example, because payment for prior utilization occurs in the subsequent year, it is discounted; however, the discount factor is not likely to be large enough to be of any quantitative importance.

A second seeming difference in incentives between use of prior and current utilization turns out on closer inspection also to be unimportant. If an individual disenrolls or dies, the organization could receive no adjustment if prior utilization is used as an adjuster, because there is no future AAPCC to be adjusted for that person, whereas the organization does receive an adjustment if current utilization is used. Thus, use of prior utilization may seem to be bad in the case of death (the organization receives no compensation for the extra expenses before death) and good in the case of disenrollment (it reduces the incentive to dump). In fact, however, the two methods can be made equal on this score. If prior utilization is used as an adjuster, an extra payment could be made after the fact for a deceased person. Similarly, if desired, payment based on current utilization could be retroactively scaled down for those who voluntarily disenroll.

Hence, we may conclude that prior and current utilization are similar in their incentive effects (or can be made so), but current utilization is a more sensitive measure of predictable variation in $\mu(it)$.

More generally, many believe that incentive problems result if any part of the payment is based on utilization, either prior or current, but one should ask if fee-for-service prices exceed marginal cost.[3] Unfortunately, given the administered price nature of insurer reimbursement and the extreme difficulty of observing marginal cost, prices exceeding

---

[2]The two proposals are not strictly comparable. The author proposes using an unspecified weighted average of current utilization (or fee-for-service) and capitation; those recommending prior utilization in effect establish a weight for it through the regression coefficient that prior utilization receives. For purposes of this discussion, assume that the weight on current utilization is about the same as for past utilization.

[3]Pauly (1980) points out that fee-for-service prices equal to marginal cost will give the physician an incentive to act as a perfect agent. The standard moral hazard, of course, remains.

marginal cost may be common. While some worry about excessive fee-for-service utilization for this reason and because the patient is insured, others worry that pure capitation will produce underutilization. If both worries are justified, a blend of the two should produce an appropriate amount of utilization.[4]

Optimal weights in a blend are a topic for future research. It is not obvious, however, that the weight on utilization should be the same for every enrollee, as the current proposals for incorporating prior utilization as an adjuster implicitly suggest. For example, in classes of enrollees with higher variances (e.g., the very old), capitation will require that the HMO accept higher risk. Perhaps in such classes one should weight the fee-for-service system payment more heavily.

It is also likely that at least some HMOs will object to anything less than complete capitation, for the same reason that the original Medicare reimbursement regulation was not popular: HMOs prefer not to bill fee-for-service. Two responses might be made:

1. Prior to PPS, hospitals were not set up to bill using DRGs. Just as hospitals have adapted to DRGs, so too could HMOs adapt to fee-for-service. Indeed, most HMOs already have a capability for generating fee-for-service bills for cases such as those covered by Workman's Compensation.

2. HMOs may argue that they would not engage in unethical behavior such as dumping, despite the monetary incentive to do so. It is questionable that all HMOs would show such restraint,[5] but one could put the argument to the test, using a blend as a fallback position.

GOALS AND MEANS

The Medicare program seeks to achieve several goals; two that are central to the present argument are economic efficiency and access for beneficiaries. Fee-for-service (with prices above marginal cost) poses no access problem, but is clearly inefficient and costly. The movement toward reform embodied in PPS suggests that taxpayers are unwilling to pay the costs of a fee-for-service system with prices above marginal costs. Capitation may stimulate greater economic efficiency,[6] but poses an access problem for high-expected-cost beneficiaries if the capitated

[4]For a similar argument in a different context see Ellis and McGuire (1986) and Ginsburg et al. (1986).

[5]Elsewhere (Newhouse 1982), the author has made the analogy with community rating and experience rating. Even if a given group of HMOs fully intend to community-rate, the actions of others may force them not to do so.

[6]The problems of setting a proper price level have been no more than touched on here. Without a proper price level, capitation will obviously not be efficient either.

rate is applied to a group with diverse expected costs. Thus, it appears difficult to satisfy both goals simultaneously. In choosing how to compromise between these two goals, a blend opens many more options than merely the two extreme points of pure fee-for-service or pure capitation.

There is one other argument for a blend: Although moving partway toward a capitated system by using a blend of capitation and fee-for-service will be seen by some as limiting the gains of a pure capitated system, it also limits the risks. At a minimum, therefore, it would seem to be a good candidate for a transition strategy.

REFERENCES

Anderson, G., and Knickman, J.: Adverse Selection Under a Voucher System: Grouping Medicare Recipients by Level of Expenditure. *Inquiry* 21:135–143, Summer 1984(b).

Anderson, G., and Knickman, J.: Patterns of Expenditure Among High Utilizers of Medical Care Services: The Experience of Medicare Beneficiaries from 1974 to 1977. *Medical Care* 22:143–149, February 1984(a).

Anderson, J., and Gertman, P.: Monitoring of HMO Enrollment Practices. AAPCC Technical Note No. 5. Health Policy Center, Heller Graduate School, Brandeis University, 1983 (cited in Lubitz, Beebe, and Riley, 1985).

Anderson, J., and Gruenberg, L.: Analyzing Alternative Methods for Establishing Medicare Capitation Payments to HMOs. Paper presented at the 111th annual meeting of the APHA, Dallas, Texas (cited in Lubitz, Beebe, and Riley, 1985).

Beebe, J., Lubitz, J., and Eggers, P.: Using Prior Utilization to Determine Payments for Medicare Enrollees in Health Maintenance Organizations. *Health Care Financing Review* 6(3):27–38, Spring 1985.

Buchanan, J., and Cretin, S.: Fee-for-Service Health Care Expenditures: Evidence of Selection Effects Among Subscribers who Choose HMOs. The RAND Corporation, R-3341-HHS, 1986.

Carter, G. M., and Ginsburg, P.: The Medicare Case Mix Index Increase: Medical Practice Changes, Aging, and DRG Creep. The RAND Corporation, R-3292-HCFA, 1985.

Carter, G. M., and Rolph, J. E.: "Empirical Bayes Methods Applied to Estimating Fire Alarm Probabilities. *Journal of the American Statistical Association* 69:880–885, December 1974.

Chassin, M. R., et al.: Variations in the Use of Medical and Surgical Services by the Medicare Population, *New England Journal of Medicine* 314:285–290, January 30, 1986.

Danzon, P.: Hospital 'Profits': The Effects of Reimbursement Policies. *Journal of Health Economics* 1:29–52, May 1982.

Detsky, A. S., et al., "Prognosis, Survival, and the Expenditure of Hospital Resources for Patients in an Intensive-Care Unit," *New England Journal of Medicine*, 305:667–672, September 17, 1981.

Efron, B., and Morris, C.: Limiting the Risk of Bayes and Empirical Bayes Estimators: Part II, The Empirical Bayes Case. *Journal of the American Statistical Association* 67:130–139, March 1972.

Efron, B., and Morris, C.: Stein's Paradox in Statistics, *Scientific American* 236:119–127, May 1977.

Eggers, P.: Risk Differential Between Medicare Beneficiaries Enrolled and Not Enrolled in an HMO. *Health Care Financing Review* 1(3):91–100, Winter 1980.

Eggers, P., and Prihoda, R.: Preenrollment Reimbursement Patterns of Medicare Beneficiaries Enrolled in At Risk HMOs. *Health Care Financing Review* 4(1):55–73, September 1982.

Ellis, R. P., and McGuire, T.: Provider Behavior Under Prospective Reimbursement: Cost Sharing and Supply. *Journal of Health Economics* 5:129–151, June 1986.

Federal Register, January 10, 1985, pp. 1378–1381.

Fisher, C.: Differences by Age Group in Health Care Spending. *Health Care Financing Review* 1(4):65–90, Spring 1980.

Garber, A., Fuchs, V. R., and Silverman, J.: Case Mix, Costs, and Outcomes—Differences Between Faculty and Community Services in a University Hospital. *New England Journal of Medicine* 310:1231–1237, 1984.

Gaus, C., Cooper, B., and Hirschman, C.: Contrasts in HMO and Fee-for-Service Performance. *Social Security Bulletin*, 3–14, May 1976.

Ginsburg, P. B., et al.: Planning a Demonstration of Per-Case Reimbursement for Inpatient Physician Services Under Medicare. The RAND Corporation, R-3378-HCFA, 1986.

Greenlick, M. R.: Comments on Medical Capitation Payments to HMOs. *Advances in Health Economics and Health Services Research*, Vol. 6, R. M. Scheffler and L. F. Rossiter (eds.). JAI Press, Greenwich, Conn., 1985.

Griffith, M. J., Baloff, N., and Spitznagel, E. L.: Utilization Patterns of Health Maintenance Organization Disenrollees. *Medical Care* 22:827–834, 1984.

Health Care Financing Review: Health Care Financing Trends. 6(1):97, Fall 1984.

Held, P. J., and Pauly, M. V.: Competition and Efficiency in the End Stage Renal Disease Program. *Journal of Health Economics* 2:95–118, August 1983.

Hennelly, V., and Boxerman, S.: Disenrollment from a Prepaid Group Plan: A Multivariate Analysis. *Medical Care* 21:1154–1167, December 1983.

Iglehart, J.: The Future of HMOs. *New England Journal of Medicine* 307:451–456, 1982.

Iglehart, J.: Medicare Turns to HMOs. *New England Journal of Medicine*, 312:132–136, 1985.

Jackson-Beeck, M., and Kleinman, J.: Evidence for Self Selection Among Health Maintenance Organization Enrollees. *Journal of the American Medical Association* 250:2826–2829, 1983.

Kunkel, S. A., and Powell, C. K.: The Adjusted Average Per Capita Cost under Risk Contracts with Providers of Health Care. *Transactions of the Society of Actuaries* 33:57–66, 1981.

Lewis, K.: Comparison of Use by Enrolled and Recently Disenrolled Populations in a Health Maintenance Organization. *Health Services Research* 19:1–22, April 1984.

Lubitz, J.: Health Status Adjustments for Medicare Capitation. Mimeo, June 4, 1985.

Lubitz, J., Beebe, J., and Riley, G.: Improving the Medicare HMO Payment Formula. *Advances in Health Economics and Health Services Research*, Vol. 6. R. M. Scheffler and L. F. Rossiter (eds.), JAI PRESS, Greenwich, Conn., 1985.

Lubitz, J., and Prihoda, R.: Use and Cost of Medicare Services in the Last 2 Years of Life. *Health Care Financing Review*, 5(3): 117–131, Spring 1984.

Luft, H. S.: *Health Maintenance Organizations.* John Wiley & Sons, Inc., New York, 1981.

McCall, N., and Wai, H. S.: An Analysis of the Use of Medicare Services by the Continuously Enrolled Aged. *Medical Care* 21:567–585, June 1983.

McClure, W.: On the Research Status of Risk-Adjusted Capitation Rates. *Inquiry* 21:205–213, Fall 1984.

Manning, W. G., et al.: A Controlled Trial of the Effect of a Prepaid Group Practice on Use of Services. *New England Journal of Medicine* 310:1505–1510, 1984.

Manning, W. G., Newhouse, J. P. and Ware, J. E., Jr.: The Status of Health in Demand Estimation. *Economic Aspects of Health.* V. R. Fuchs (ed.), University of Chicago Press, 1982.

Mechanic, D., Weiss, N., and Cleary, P.: The Growth of HMOs: Issues of Enrollment and Disenrollment. *Medical Care* 21:338–347, March 1983.

Morrisey, M., et al.: Hospital Rate Review: A Theory and Empirical Review. *Journal of Health Economics* 3:25–47, April 1984.

Newhouse, J. P.: Cream Skimming, Asymmetric Information, and a Competitive Insurance Market. *Journal of Health Economics* 3:97–100, April 1984.

Newhouse, J. P.: Is Competition the Answer? *Journal of Health Economics* 1:109–115, May 1982.

Nobrega, F., et al.: Hospital Use in a Fee-for-Service System. *Journal of the American Medical Association* 247:806–810, 1982.

Pauly, M. V.: Doctors and Their Workshops: Economic Models of Physician Behavior. University of Chicago Press, 1980.

Pauly, M. V.: Is Cream-Skimming a Problem for the Competitive Medical Market? *Journal of Health Economics* 3:87–95, April 1984.

Pauly, M. V.: What is Adverse About Adverse Selection? *Advances in Health Economics and Health Services Research.* R. M. Scheffler and L. F. Rossiter (eds.), JAI Press, Greenwich, Conn., 1985.

Rothschild, M., and Stiglitz, J.: Equilibrium in Competitive Insurance Markets: An Essay in the Economics of Imperfect Information. *Quarterly Journal of Economics* 40:629–649, November 1976.

Thomas, J. W., et al.: Increasing Medicare Enrollment in HMOs: The Need for Capitation Rates Adjusted for Health Status. *Inquiry* 20:227–239, Fall 1983.

Thomas, J. W., and Lichtenstein, R.: Functional Health Measure for Adjusting Health Maintenance Organization Capitation Rates. Health Care Review 7(3):85–95, Spring 1986.

Trieger, S., Galblum, T. W., and Riley, G.: HMOs: Issues and Alternatives for Medicare and Medicaid. General Accounting Office, No. 522–527/6037 1–3. Washington, D.C., 1982.

Welch, W. P.: Medicare Capitation Payments to HMOs in Light of Regression Toward the Mean in Health Care Costs. *Advances in Health Economics and Health Services Research.* R. M. Scheffler and L. F. Rossiter (eds.), JAI Press, Greenwich, Conn., 1985(a).

Welch, W. P.: Regression Toward the Mean in Medical Care Costs. *Medical Care* 23:1234–1241, November 1985(b).

Wennberg, J., and Gittelsohn, A.: Small Area Variations in Health Care Delivery. *Science* 182:1102–1108, December 14, 1973.

Ware, J. E., Jr., et al.: Comparison of Health Outcomes at a Health Maintenance Organization with Those of Fee-for-Service Care. *The Lancet* 313:1017–1022, May 3, 1986.

Wersinger, R. P. and Sorenson, A. A.: Demographic Characteristics and Prior Utilization Experience of HMO Disenrollees Compared with Total Membership. *Medical Care* 20:1188–1196, December 1982.

Wollstadt, L. J., Shapiro, S., and Bice, T. W.: Disenrollment from a Prepaid Group Practice: An Actuarial and Demographic Description. *Inquiry* 15:142–150, June 1978.

RAND/R-3455-HCFA

CMS LIBRARY
3 8095 00014195 8